Genetic Analysis of Mountain Pine Beetle Populations

Final Report #: 03-JV-11222033-314

Library

Final Report:
Genetic Analysis of Mountain Pine Beetle Populations

Supported in part by funds provided by the Rocky Mountain Research Station,
U. S. Forest Service, U. S. Department of Agriculture
FS Agreement Number 03-JV-11222033 - 314
May 31, 2004

from:
Karen Mock
Forest, Range, and Wildlife Sciences Department
Utah State University

to:
Barbara Bentz
USDA Forestry Sciences Laboratory

Introduction:

The mountain pine beetle (MPB) (*Dendroctonus ponderosae* Hopkins) is an ecologically and economically important pest species associated with pine forests in western North America (Safranyik et al. 1981; Amman and Cole 1983). Although this insect occurs from Baja California to southern British Columbia, its range is more restricted than that of its potential host trees, suggesting that it may be limited by factors such as climate, competition, and regional host tree defenses, in addition to host tree distribution (Safranyik 1978; Cates & Alexander 1982; Ayers & Lombardero 2000; Logan & Powell 2001). Genetic structuring among populations of the MPB is not well understood, and to date has only been addressed in studies of limited geographic scope. Stock et al. (1984) found only low levels of variation among 15 populations of MPB sampled in seven western states using a panel of allozymes. Sturgeon & Mitton (1986), using more highly polymorphic allozymes, found genetic differentiation among local populations and among different tree hosts in north central Colorado. Langor and Spence (1991) also found pronounced differentiation among local populations in Alberta and British Columbia, but little evidence for genetic subdivision among sympatric host tree types.

The broad objective of our current collaborative work with MPB is to assess rangewide genetic structure using a highly polymorphic marker system, amplified fragment length polymorphism (AFLP), in order to provide a framework for studies on adaptive variation and evolution in this insect. AFLP is a dominant, multilocus, predominantly neutral marker system that can be extremely useful in describing landscape-scale population structure with little preknowledge of genomic structure. The goals of this Joint Venture Agreement were to 1) enable protocol optimization (collection, preservation, extraction, and AFLP analysis) in preparation for a large-scale study, 2) to assemble a large set of extracted DNA samples for this study from field collections, and 3) to demonstrate the utility of these protocols by assembling a preliminary dataset describing rangewide genetic structuring MPB.

Protocol Optimization:

Genetic analysis on MPB requires the collection and extraction of high quality DNA from individual beetles. The quality of DNA can be negatively impacted by improper preservation and storage following live collection or by the co-extraction of pigments or DNA from other organisms that interfere with downstream analyses.

Collection & Storage protocols: In order to assess preservation and storage protocols, experiments were conducted with drying, freezing following live collection, collection directly into a lysis buffer, collection directly into 95% ethanol, and a combination of collection into ethanol or lysis buffer followed by storage in a freezer. Our results suggest that collection of live insects followed by freezing at -80°C yielded the highest quality DNA. However, this approach is not always feasible in field, and we found that collection directly into 95% ethanol, followed by freezing upon arrival in the lab, is the optimal field protocol to preserve DNA quality when live collection & freezing are not practical.

Dissection for DNA extraction: Because MPB are know to harbor mycangial fungi, and because these fungi may interfere with downstream genetic analyses (particularly amplified length polymorphism, or AFLP analysis) (Vos et al. 1995), we assessed the effect of decapitating the beetles prior to DNA extraction. In addition, we assessed the impact of dissecting out only the thorax for extraction, in order to avoid potential contamination from gut symbionts. Our results from extraction and AFLP analyses indicated that while decapitation does reduce the overall yield of DNA from individual beetles, the resulting quality and quantity of DNA are sufficient for downstream analyses. Dissection of the thorax, however, did reduce the total yield to the point where replicate analyses would be limited. AFLP analysis of these sample sets indicated that the inclusion of abdomens did not yield scorable spurious bands in AFLP analysis that would bias the data, suggesting that DNA contamination from gut endosymbionts may not be a major concern. However, we did refine our protocol to include decapitation prior to extraction and inclusion of separate thorax vs. abdomen extractions for 5% of the samples in AFLP analysis as quality assurance measures.

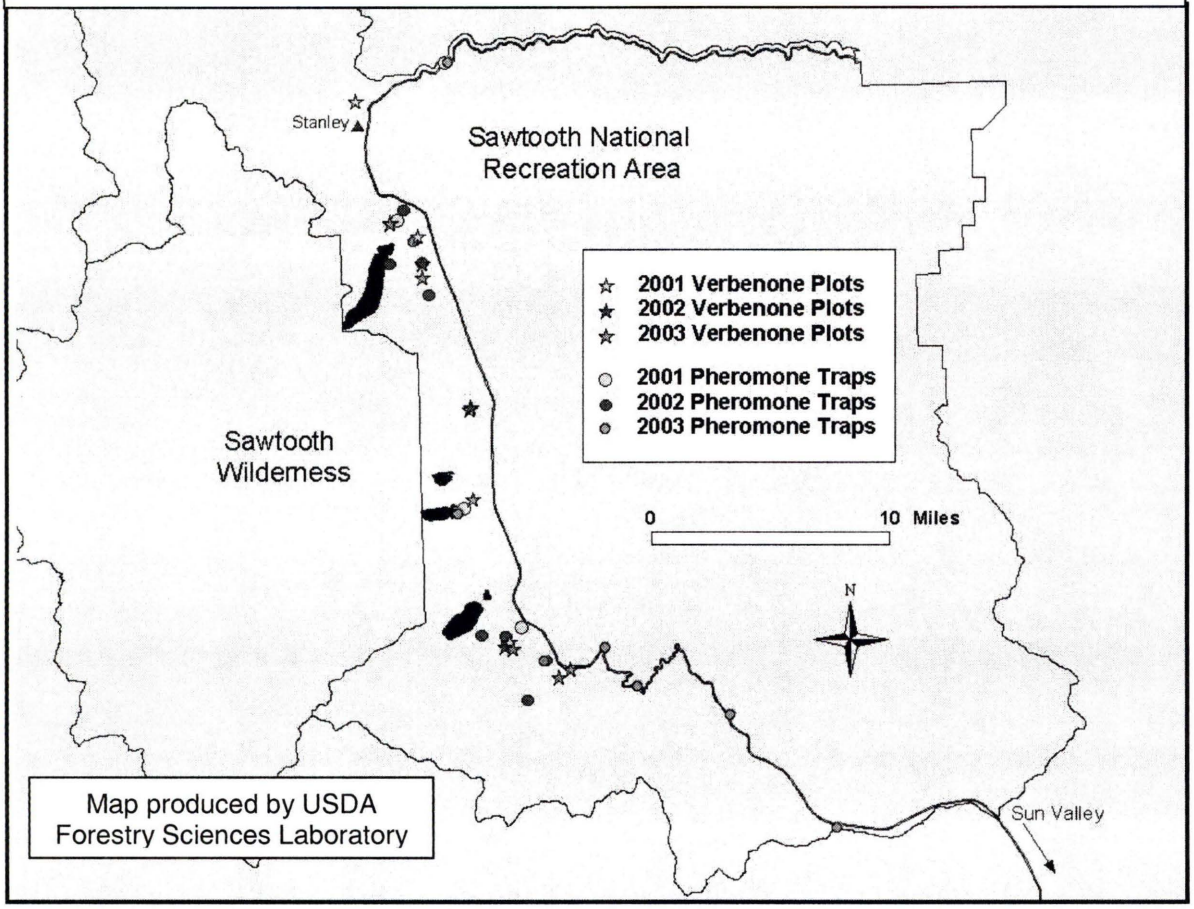
DNA extraction protocols: We assessed the DNA quality and quantity resulting from a variety of extraction protocols, including salt/chloroform extraction, salting-out extraction, and Eppendorf DNA extraction spin columns. Overall we found that organic (salt/chloroform) extractions provided the best results, and yielded DNA that was not contaminated with PCR-inhibiting pigments.

DNA Dataset Compilation:

Sawtooth National Recreation Area (SNRA) MPB sample collection & DNA extraction: MPB Samples were collected by Forest Service employees from 8 pheromone traps on a weekly basis for 14 weeks over the emergence period in the SNRA. The 8 traps represented both a geographic and elevational gradient (Figures 1 & 2). Trap collections from all locations over this period indicate a bimodal pattern of emergence (Figure 3). Up to 40 individuals per week per site were preserved in 95% ethanol and are stored at -80°C. Our initial DNA extraction efforts were designed to assess population-level differences along both a spatial and temporal gradient. Four locations (Petit, River, Galena Summit, and Boulder) were chosen to represent the spatial gradient, and 6 weeks (3,

5, 7, 8, 11, and 14) were chosen to represent the temporal gradient. With a target of 22 samples per site per week, this is a total of 528 samples. To date, 181 of these have been extracted (Table 1). Extractions on this sample set will continue in the summer and fall 2005 using alternative (non-RMRS) funding sources.

Figure 1. MPB collection sites in SNRA, summer 2003. Samples were collected at 2003 pheromone trap sites. Map produced by USDA RMRS.



Western regional MPB sample collection, DNA extraction, and AFLP analysis: A minimum of 20 individuals from each of 9 regions across the MPB range (San Bernardino NF, Uinta NF, Wallowa-Whitman NF, Coconino NF, Helena National Forest (NF), Klamath NF, Idaho Panhandle NF, Sawtooth NRA, and BC Provincial Lands, Canada) (Figure 4)(Table 2) were collected by RMRS personnel in 2003 and 2004 for genetic analysis. DNA was extracted from a subset of individuals from each population using a salt-chloroform method (Mullenbach et al. 1989). The quality and quantity of the extracted DNA was assessed on 0.7% agarose gels. In order to characterize nuclear divergence and diversity among these populations, Amplified Fragment Length Polymorphism (AFLP) marker profiles were generated following a modified version of basic procedures from Vos et al. (1995), using ten selective primer combinations: EcoRI-AGG and MseI-AC, EcoRI-AGG and MseI-AG, EcoRI-ACG and MseI-AC, EcoRI-ACG and MseI-AG, EcoRI-AGC and MseI-AC, EcoRI-ACC and MseI-AC, EcoRI-ATC and MseI-AC, EcoRI-AGC and MseI-AG, EcoRI-ACC

Figure 2. SNRA MPB collection sites for Summer 2003, represented by distance and elevation. Figure produced by USDA RMRS.

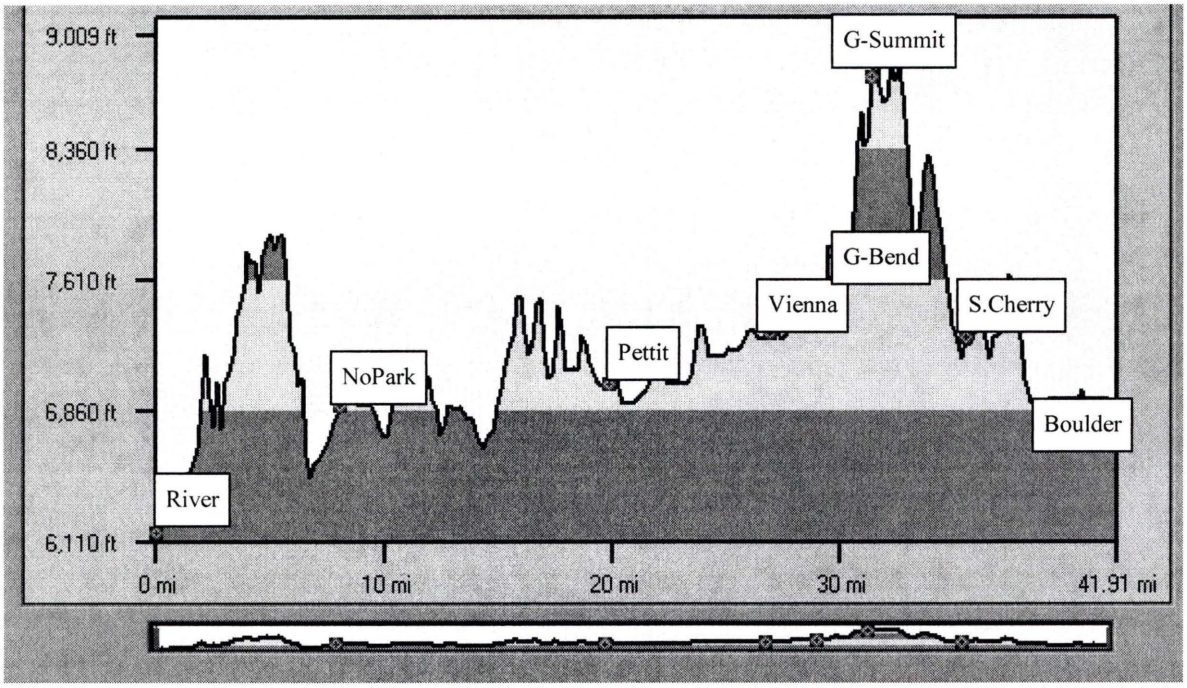
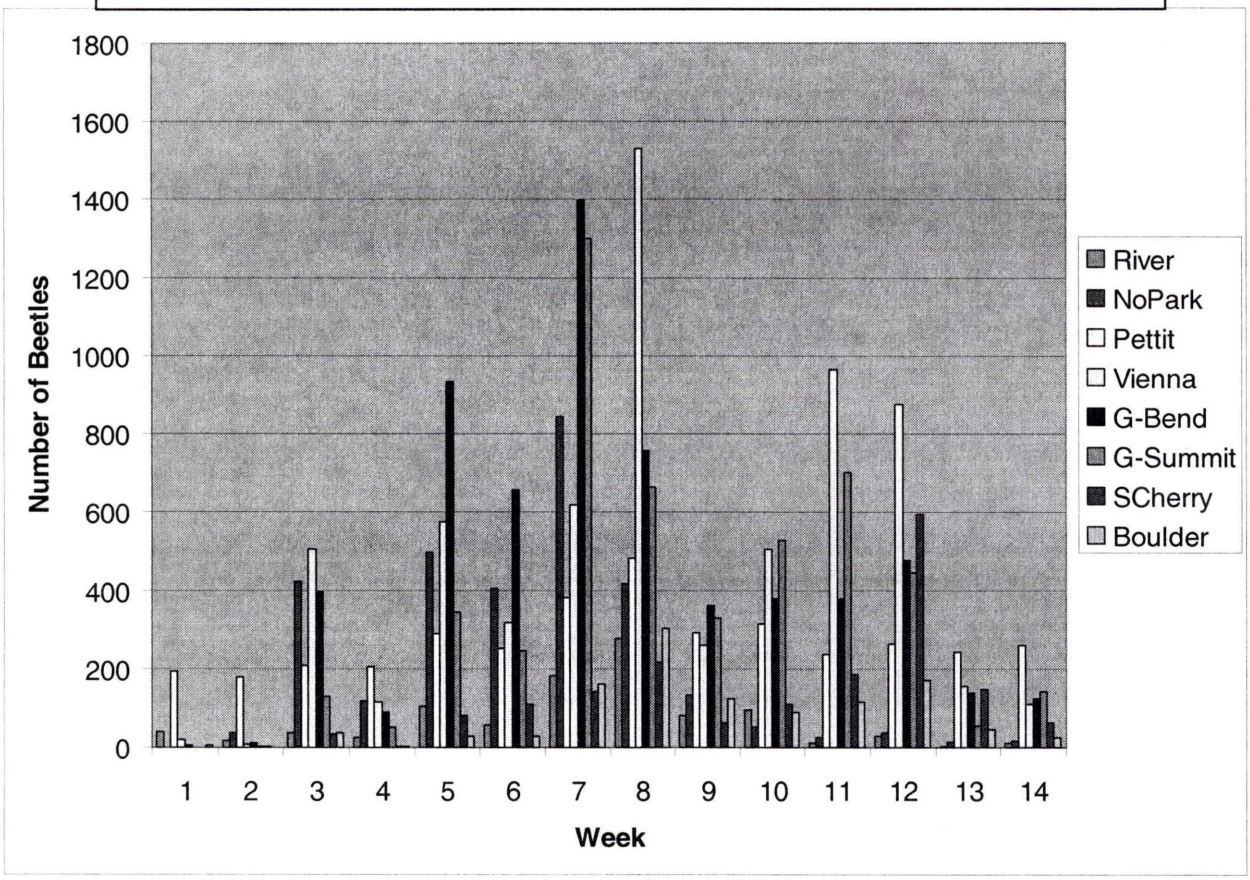


Figure 3. Temporal trap rate data associated with SNRA MPB collections, summer 2003.



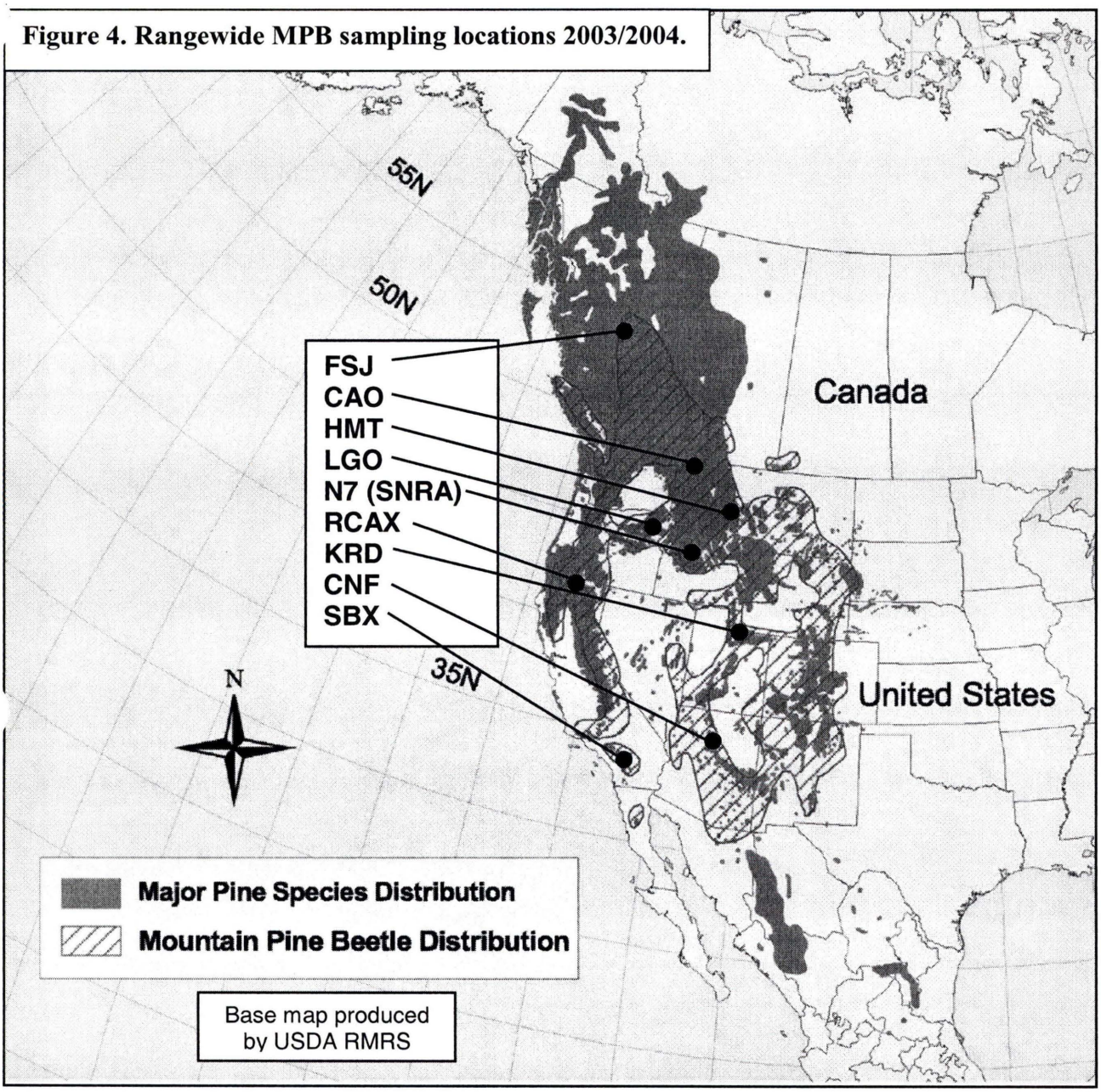


Figure 4. Rangewide MPB sampling locations 2003/2004.

and MseI-AG, and EcoRI-ATC and MseI-AG. Resulting amplicons were run on a sequencing gel with a ROX 400 (ABI) size standard using an ABI 3100 automated sequencer. Genographer 1.6 software (Benham 2001) was used to visualize and score the gel image. To date, two of the selective primer combinations have been completed for all samples in the dataset (Table 2), providing the basis for a limited phylogeographic assessment of MPB. Generation of AFLP data for the remaining samples is continuing through summer 2004 with funding from other sources. AFLP markers are scored if they were polymorphic across the dataset (95% criterion) and if they can be scored unambiguously. Scoring is performed without reference to sample or population identity.

Table 1. DNA extractions performed to date from sample locations in the Sawtooth National Recreation Area (SNRA).		
Site	**Collection Week**	**Extractions**
River	7	22
NoPark (N7)	**	22
Petit	5	21
Petit	7	22
G-Summit	5	22
G-Summit	7	22
Boulder	5	6
Boulder	7	22
Boulder	11	22

Table 2. Western regional MPB populations included in preliminary AFLP genetic analysis.				
Population	**Location**	**Date Collected**	**Tree Host**	**N (AFLP)**
SBS	Arrowhead District Office, San Bernardino NF, CA	27-Jan	Sugar Pine	19
SBL	Mt. San Jacinto State Park, CA	29-Jan	Lodgepole Pine	19
SBP	Arrowhead Lake, San Bernardino NF, CA		Ponderosa Pine	0
KRD	Uinta NF	7-Aug-03		22
LGO	Moss Springs, Wallowa-Whitman NF, OR	7-Aug-03	Lodgepole Pine	21
CNF	San Francisco Peaks, Coconino NF, AZ	15-Aug-03	Limber Pine	13
HMT	Helena NF, MT	5-Aug-03	Ponderosa Pine	8
RCAW	Goosenest Mountain, Klamath NF, CA	29-Jul-03	Whitebark Pine	18
RCAL	Martin's Dairy, Klamath NF, CA	24-Jul-03	Lodgepole Pine	40
FSJ	John Prince Research Forest, Fort St. James, BC, Canada			22
N7	Sawtooth National Recreation Area, ID	Summer 2003	Lodgepole Pine	14
CAO	Fisher Peak, Idaho Panhandle NF, Coeur d'Alene ID	12-Aug-03	Whitebark Pine	18

Preliminary phylogeographic analysis of MPB using AFLP markers.

AFLP Methodology. Using the protocols described above, 34 AFLP loci were found to be polymorphic and reliably scorable over all samples (Table 2) using the first two selective primer combinations. Seven of these markers were found to be associated with sex (present exclusively or nearly exclusively in males), and were removed from the dataset. These sex-associated markers, however, will be used in the future to reconstruct male-specific phylogenies. Forty one (19%) of the 214 samples (including representatives from all populations) were replicated following DNA extraction to provide a conservative assessment of the error associated with extraction, PCR and scoring.

Data Analysis. AFLP profiles over 27 polymorphic loci were assembled for all populations (Table 2). Because several of these populations represented host tree differences in sympatric locations (RCAL & RCAW; SBL & SBS), we first performed exact testing (Raymond and Rousset 1995) to determine whether these sympatric populations were genetically distinct. Allele frequencies were estimated using the Taylor expansion method (Lynch and Milligan 1994) under the assumption of Hardy-Weinberg genotypic equilibrium (HWE) using Tools for Population Genetic Analysis (TFPGA) software (Miller 1997). The assumption of HWE is required because AFLP markers are dominant (i.e. homozygous and heterozygous genotypes cannot be distinguished). Exact testing among all pairs of sympatric populations was performed using the Markov Chain Monte Carlo approach provided in TFPGA. The results of these analyses were used to determine whether these sympatric populations were combined in subsequent analysis. Exact testing was performed again on the redefined populations, as described above, using a Bonferroni-corrected p-value. Population differentiation was also assessed using analysis of molecular variation (AMOVA) with Arlequin software (Schneider et al. 2000). Statistical significance of the F_{ST} estimates was determined using 1000 permutations of AFLP profiles among populations to estimate a null distribution. A UPGMA cluster analysis of populations, based on Nei's (1978) genetic distance was constructed, and the strength of nodes was assessed by bootstrapping 1000 times over loci, using TFPGA. A simple matching distance matrix of individuals was also constructed and used to perform principal coordinates analyses (PCoA) in NTSYS (Rohlf 2002), in order to visualize the strength and clustering of individuals among populations. Population-level diversity was assessed by estimating unbiased heterozygosity and percent polymorphic loci (95% criterion) using TFPGA software.

Results. Based on the replicate set described above, the overall extraction and AFLP error rate was estimated at 1.03%. Exact testing of sympatric populations indicated that differentiation was not detected between RCAL and RCAW ($p = 0.8782$) or between SBS and SBL ($p = 0.5963$). These pairs of populations were combined (designated RCAX and SBX) for subsequent analyses. Exact testing indicated significant differentiation (Bonferroni-corrected $p < 0.0014$) among all pairs of populations except HMT and CAO ($p = 0.8218$), CAO and N7 ($p = 0.8564$), LGO and HMT ($p = 0.0107$), N7 and HMT ($p = 0.7698$), and LGO and N7 ($p = 0.0270$). Differentiation between LGO and CAO approached significance ($p = 0.0004$). Overall, population differentiation was pronounced, with an F_{ST} estimation of 0.304 ($p < 0.00001$). F_{ST} estimates indicated that all pairs of populations were significantly differentiated except N7 and CAO, HMT and CAO, and N7 and

Table 3. Matrix of AMOVA results for all pairwise populations of mountain pine beetle. Non-significant ($p > 0.0015$) F_{ST} estimates are shown in bold italics.

	CAO	CNF	FSJ	HMT	KRD	LGO	N7	RCAX	SBX
CAO	0.00000								
CNF	0.45707	0.00000							
FSJ	0.28187	0.58805	0.00000						
HMT	***0.05912***	0.51749	0.40131	0.00000					
KRD	0.24683	0.33002	0.42314	0.31626	0.00000				
LGO	0.16020	0.54235	0.43791	0.21927	0.38612	0.00000			
N7	***0.03972***	0.49990	0.31522	***0.08217***	0.26646	0.14899	0.00000		
RCAX	0.14664	0.38553	0.35577	0.21229	0.29675	0.09487	0.17919	0.00000	
SBX	0.27716	0.45545	0.48097	0.28249	0.44198	0.29808	0.32554	0.14773	0.00000

Figure 5. UPGMA clustering dendrogram of MPB populations. Numbers at nodes represent the proportion of 1000 bootstrap replicates over loci

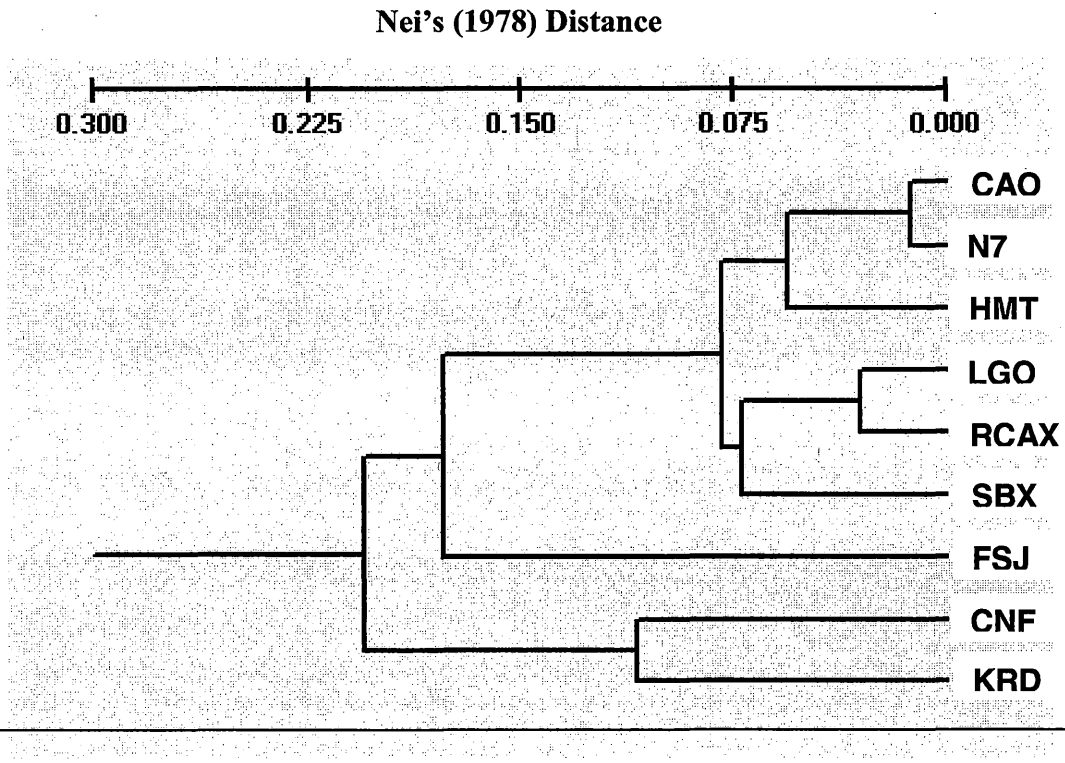
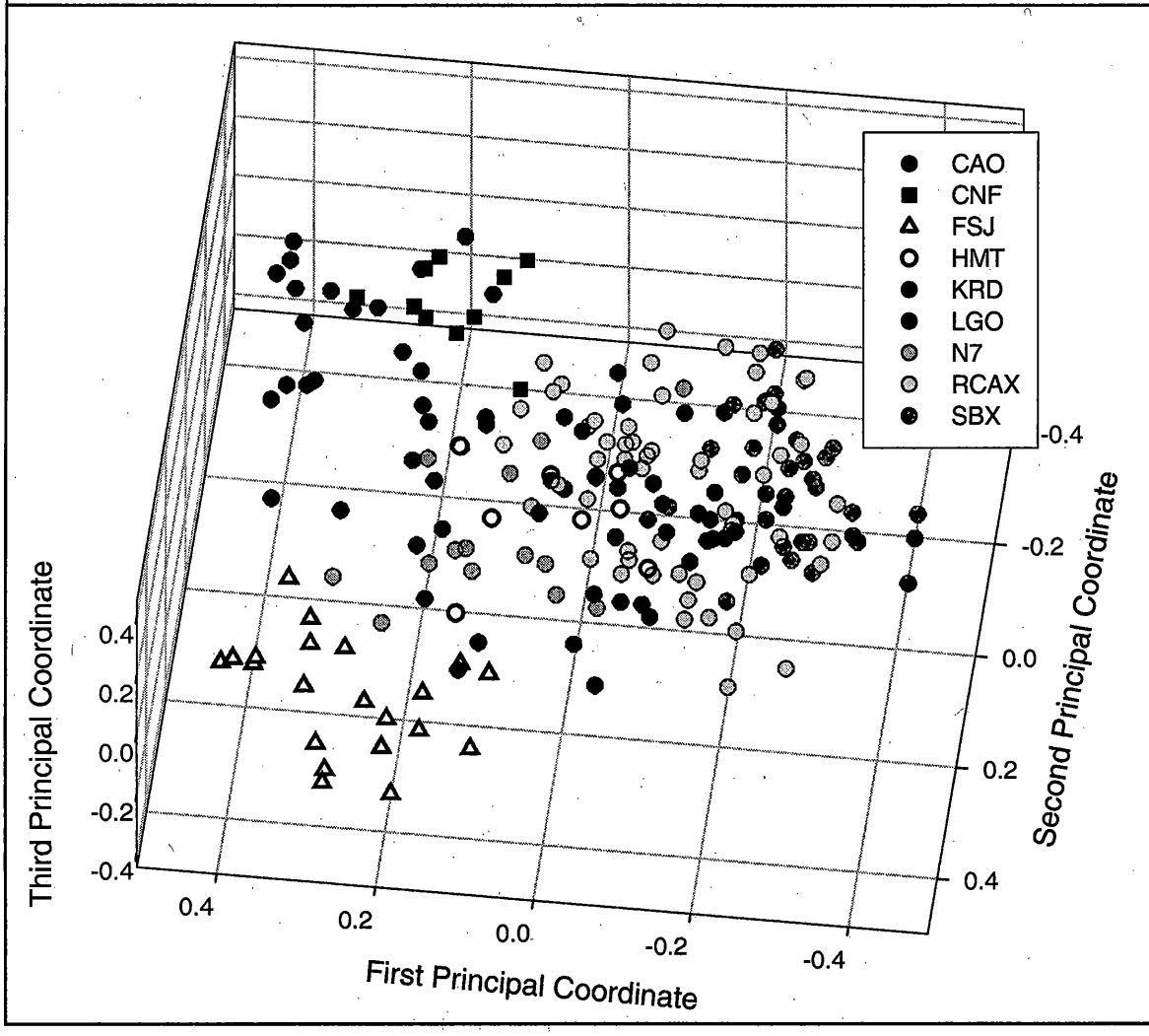


Figure 6. Principal coordinates plot of MPB individuals, designated by population, based on a simple matching distance matrix constructed from AFLP data.



HMT (Table 3), corroborating the results of exact testing. A UPGMA dendrogram of populations (Figure 5) also suggested a close alliance between CAO, N7, and HMT, albeit with weak bootstrap support. FSJ, CNF, and CRD appeared to be differentiated from the other populations in the UPGMA dendrogram, with CNF and KRD allied (Figure 5). This pattern of similarity suggests an isolation-by-distance distribution of genetic variation across the range of the MPB. Clustering of individuals, based on a genetic distance matrix, is illustrated in a principal coordinates plot (Figure 6), which also indicates an isolation-by-distance pattern of genetic divergence among populations. Measures of diversity were similar among populations, with the exception of CNF, which appears to have relatively low diversity (Table 4).

Table 4. Population-specific diversity indices in MPB. Unbiased heterozygosity (H) and percent polymorphic loci (%P)(95% criterion) are provided.

Population	H	%P
CAO	0.28	74.07
CNF	0.17	44.44
FSJ	0.20	51.85
HMT	0.25	66.67
KRD	0.29	62.96
LGO	0.27	70.37
N7	0.28	70.37
RCAX	0.29	88.89
SBX	0.20	62.96

Discussion. The populations (LGO, HMT, N7 and CAO) are more geographically proximal than other populations in the dataset, and they do not appear to be consistently differentiated from one another, suggesting high levels of gene flow. There appear to be lower levels of gene flow between this group of populations and FSJ, CNF, and KRD. CNF and KRD are genetically distinct populations, but they appear to be more similar to each other than they are to other populations. The alliance of CNF and KRD indicates that MPB in the Uinta Mountains share more gene flow with the MPB on the Colorado Plateau than with the MPB in central Idaho. The apparent discontinuity between northern Utah and southern Idaho may be due to a constriction in the range of host trees, historical patterns of MPB range expansion and contraction, or some kind of an ecological barrier. The SBX population was more closely allied with the populations in northern California, Oregon, Idaho, and Montana than with the geographically more proximal CNF population. This pattern is likely due to the effectiveness of the Mojave Desert as a barrier to gene flow. The FSJ population was also distinct from other populations, a pattern that could be due to isolation by distance. These patterns of genetic divergence suggest that the MPB is driven by a combination of dispersal abilities and host tree occurrence.

Recommendations. Based on a limited dataset of 27 AFLP loci, we were able to detect a significant level of population structuring in the MPB. We predict that this pattern will become more pronounced as additional loci are analyzed and as population sample sizes are increased. Additional sample populations in intermediate regions would allow us to distinguish between a broad pattern of isolation by distance and more abrupt barriers to gene flow among populations. This information could be extremely useful in managing MPB outbreaks and migration patterns. AFLP loci are generally considered to be an indication of neutral processes. However, as additional AFLP loci are scored across the dataset, we are increasingly likely to encounter loci that are under differential selection pressures across the large range of this insect and do not fit the neutral expectation (Luikart et al. 2003). These loci will be excellent candidates for the identification of ecologically important genes. We recommend:

- the continued analysis of this dataset, including additional loci and individual samples,
- the development of microsatellite markers in MPB to provide a codominant, highly polymorphic marker system to investigate fine-scale gene flow patterns among regions, and
- the inclusion of additional population locations.

Project Benefits.

This Joint Venture Agreement between the RMRS and Utah State University has enabled:

a) the initiation of a new, potentially highly productive collaborative relationship between researchers at these institutions,
b) the synthesis of a preliminary dataset that is an important component of a National Science Foundation proposal (submitted by USU and RMRS January 2004, status pending),
c) the collection, archiving, DNA extraction, and initial AFLP analysis for a large dataset that will become the basis for a major research effort over the next several years.

References:

Amman G. D. and W. E. Cole. 1983. Mountain pine beetle dynamics in lodgepole pine forests part II: Population dynamics. U. S. D. A., Forest Service Intermountain Forest and Range Experiment Station, General Technical Report INT-145. Ogden, UT. 59 pp.

Ayers, M. P. and M. J. Lombardero. 2000. Assessing the consequences of global change for forest disturbance from herbivores and pathogens. Sci. Total Environ. 262:263-286.

Benham, J. J. 2001. Genographer Software, Version 1.6.0. http://hordeum.oscs.montana.edu/genographer.

Cates, R. G. and H. Alexander. 1982. Host resistance and susceptibility. Pp. 212-263 in Mitton, J. B. and K. B. Sturgeon (eds.), Bark Beetles in North American Conifers: A System for the Study of Evolutionary Biology. University of Texas Press, Austin, TX.

Langor, D. W. and J. R. Spence. 1991. Host effects on allozyme and morphological variation of the mountain pine beetle, *Dendroctonus ponderosae* Hopkins (Coleoptera: Scolytidae). The Canadian Entomologist 123: 395-410.

Logan, J. A. and J. A. Powell. 2001. Ghost forests, global warming, and the mountain pine beetle (Coleoptera: Scolytidae). American Entomologist 47(3):160-173.

Luikart, G., P. R. England, D. Tallman, S. Jordan, and P. Taberlet. 2003. Population genomics: from genotyping to genome typing. Nature Reviews Genetics 4:981.

Lynch, S. M. and G. G. Milligan. 1994. Analysis of population genetic structure with RAPD markers. Molecular Ecology 3: 91-99.

Miller, M. P. 1997. Tools for population genetic analysis (TFPGA), 1.3: a Windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by author (http://bioweb.usu.edu/mpmbio).

Mullenbach, R., P. L. Lagoda, and C. Welter. 1989. An efficient salt-chloroform extraction of DNA from blood and tissues. Trends in Genetics 5:391.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583-590.

Raymond, M. L. and F. Rousset. 1995. An exact test for population differentiation. Evolution 49: 1280-1283.

Rohlf, F. J. 2002. *NTSYSpc: Numerical Taxonomy System*, version 2.10t. Exeter Publishing, Setauket, New York.

Safranyik, L. 1978. Effects of climate and weather on mountain pine beetle populations. In Theory and Practice of Mountain Pine Beetle Management, A. A. Berryman, G. D. Amman, and R. W. Stark (eds.), pp. 77-86. Forest, Wildlife and Range Experiment Station, University of Idaho, Moscow, ID.

Safranyik, L., G. A. van Sickle, and G. H. Manning (Compilers). 1981. Position paper on mountain pine beetle problems with special reference to the Rocky Mountain Parks Region. Canadian Forestry Service, Pacific Forestry Centre, Victoria, B. C. 27 pp.

Schneider, S., Roessli, D., and Excoffier, L. 2000. Arlequin: A software for population genetics data analysis. Ver 2.001. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva.

Stock, M. W., G. D. Amman, and P. K. Higby. 1984. Genetic variation among mountain pine beetle (*Dendroctonus ponderosae*)(Coleoptera: Scolytidae) populations from seven western states. Annals of the Entomological Society of America. 77(6):760-764.

Sturgeon, K. B. and J. B. Mitton. 1986. Allozyme and morphological differentiatioin of mountain pine beetles *Dendroctonus ponderosae* Hopkins (Coleoptera: Scolytidae) associated with host tree. Evolution 40(2): 290-302.

Vos, P., R. Hogers, M. Bleeker, et al. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23:4407-4414.